DOCUMENT RESUME

ED 049 302 TM 000 491

AUTHOR	Schmalgeneier, William L.; Watson, Richard P.
TITLE	Michigan Schools: The Organization and Management of Their Testing Programs, 1970.
INSTITUTION	Michigan Univ., Ann Arbor. Bureau of School Services.
PUB DATE	70
NOTE	51p.
EDRS PRICE	EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS	Elementary Schools, Evaluation, *Information Dissemination, Secondary Schools, *Standardized Tests, State Programs, *State Surveys, Test Construction, *Testing Programs, *Test Results, Test Selection

ABSTRACT

This is the third in a series of investigations, conducted at five-year intervals, into the testing programs of Michigan school districts. The report opens with general data on testing programs and practices in the form of a tabulation of responses to a survey instrument completed by 84% of the districts that operate a K-12 program. A more specific look at operational content follows, with attention directed to the tests given, how they are administered and scored, availability of results, and the norms used. The number of districts using a specific test, the frequency of its use, and the situation in which it is used are then reviewed. Finally, the report attempts to assess the causal nature of some of the testing practices by reviewing certain information in conjunction with other information. By, for example, relating the uses of test data to the ways in which teachers learn about the data. Responses are reported variously as raw values, percentages, or weighted values, the identity being specified at each reference point. (Author/CK)

ED049302

MICHIGAN SCHOOLS:

THE ORGANIZATION & MANAGEMENT OF THEIR TESTING PROGRAMS.

1970

164

MICHIGAN SCHOOL TESTING SERVICE
THE BUREAU OF SCHOOL SERVICES
THE UNIVERSITY OF MICHIGAN

ED049302

MICHIGAN SCHOOLS: THE ORGANIZATION AND MANAGEMENT OF THEIR TESTING PROGRAMS 1970

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

A study conducted by the Michigan School Testing Service, Bureau of School Services, The University of Michigan

Prepared by William L. Schmalgemeier and Richard P. Watson

The University of Michigan
Bureau of School Services
Ann Arbor, Michigan

TM 000 491

FOREWORD

In 1958-59, the Michigan School Testing Service division of The University of Michigan's Bureau of School Services co-sponsored, with the Michigan Association of Secondary School Principals, a comprehensive study inquiring into the nature of K-12 school testing programs throughout the State of Michigan. This was followed in 1963-64 by a similar comprehensive testing survey.

As a result of requests from directors of school testing programs, the Bureau's Michigan School Testing Service is pleased to announce a third study, the results of which are summarized in this booklet.

Preparing reports of this nature is but one of several services offered to school districts by the Michigan School Testing Service. The main purpose of the Bureau's testing service is to provide administrators, guidance and testing directors, counselors, and teachers with information that will be of help in making their testing programs more meaningful and helpful.

Special recognition for the preparation of this booklet should be given to Richard Watson, acting director of the Michigan School Testing Service, and William Schmalgemeier, advisory associate to Dr. Watson.

We shall appreciate any comments or suggestions that will enable us to be of greater service.

Kent W. Leach, Director
Bureau of School Services

ACKNOWLEDGMENTS

The cooperation of the respondents to the survey has made this report possible. This is not inconsistent with the general cooperation the Michigan School Testing Service regularly experiences in working with its colleagues in local school districts. We only hope their individual efforts will be rewarded by the information contained in this report.

The example of Frank B. Womer in carrying out the previous studies set a standard which we are yet striving to attain. The counsel and advice of LeVerne S. Collet were the seeds which bore fruit in the final several chapters of this report for which we give our sincere thanks. Mention must be made of the many colleagues at the Bureau of School Services whose often expressed interest helped brighten otherwise dull days.

Special acknowledgment must go to Yvonne Gillies and Karen Reppuhn who have worked diligently coding, de-coding, typing, re-typing, and generally preparing this document for distribution. The computer facilities of The University of Michigan must also be thanked.

W.L.S.
R.P.W.

TABLE OF CONTENTS

Chapter		Page
	Foreword	*iii*
	Acknowledgements	*v*
I	The Study	3
II	The Data	5
III	The Second Study: Testing Situations	19
IV	Testing Patterns	31
V	Test Information Usage	35
VI	Test Usage: District Characteristics	41
	Conclusion	49

Chapter I

THE STUDY

This study is the third in a series of investigations into the testing programs of Michigan school districts. The two previous studies have reported to Michigan educators the testing practices in the state. Reactions received to publishing these findings seemed to dictate that the Michigan School Testing Service of The University of Michigan again survey school districts to report their testing behavior. It would appear as though an approximate five-year schedule for this survey is in the making.

The 1963-64 report emphasized strongly the changes that had occurred in the interval between it and its 1958-59 predecessor. The present document, however, will not do that. Rather, efforts will be made to describe some of the apparent interrelationships between certain pieces of reported information. In this sense the direction of the present report is more a prescription for testing use than a documentation of past performance.

This report is organized into four parts. The first part, Chapter II, is essentially a tabulation, with commentary, of the frequencies of choice to the survey instrument, item-by-item. This will satisfy the interests of those readers who wish to know what the testing practices are throughout the state.

The second part, Chapter III, cuts deeper into operational content. To gather this information a sample of one-third of the total districts was used. This was accomplished by randomly mailing a more comprehensive questionnaire to every third district on the mailing list. This was judged to be adequate to reflect the substance of the information sought.

The next part, Chapter IV, deals with the specific tests used by the various districts throughout the state. Both the number of districts using the tests and frequency of their use are reported for the more often used tests.

The final part, Chapters V and VI, is relational in nature. In it certain information is shown in conjunction with other information, the attempt being to assess the causal nature of some of the testing practices. Thus, for example, the uses to which test data are put will be reviewed in relation to the ways in which teachers learn about these data.

Two major developments have occurred since the last survey was made that have had some impact on the nature of testing programs in Michigan. While these developments are not directly reflected in the information of this report because of the nature of the questions asked, the reader should be cognizant of their presence. The first is the nearly total eradication of federally-sponsored machinations that influence psychometric practices. This has presumably enabled local option to return increasingly to testing program determination, while the concurrent loss in preferred guidelines may have been felt as well. Testing programs should be more self-justified under these circumstances, and thus more closely related to the needs of the specific district.

The other occurrence has been the 1969 Michigan legislation requiring school districts to apply common criterion measures to determine local performance. While not directly affecting the present purposes for giving standardized tests in the districts, the presence of these state wide tests and their consequent data cannot help but affect the philosophy of testing if not its function.

The population of this survey was defined as all school districts, public and private, in the state of Michigan that operate a K-12 program. The questionnaires were sent to Directors of Testing where another title was not available. Three follow-ups were made by mail but the vigor of pursuit of previous studies was not employed, the 84 percent response of total districts being adjudged as adequate to represent the practices of Michigan schools.

Results are reported variously as raw values, percentages or weighted values, the identity being specified at each reference point.

Chapter II

THE DATA

The Presence of an Organized Testing Program

Eighty-eight percent of the reporting districts indicated they have an organized program, i.e., schedule, of testing. This may mean anything in terms of planning from merely having the plan committed to paper to having a balanced, purposeful means of gathering information. But "organized" it was for 88 percent.

Organized programs ranged from obviously brand new activities to the veterans of fifteen years or more. The median age is approximately nine years. Table I, which shows the duration of organized programs, also indicates that almost two-thirds of the districts reported testing schedules which could not have been included in the first of these surveys.

All those districts not reporting any structured program gave tests, with the exception of two districts, but the information suggested that independent decisions by a principal and/or teacher determined what was given and when. Table II shows the variety of patterns of the school districts which reported no organized program.

Management of the Testing Program

As the responsibilities for the maintenance of the various activities of a school district are assigned, one that becomes increasingly necessary is the management of the school testing program. Too great a price is paid by the district that permits spontaneous testing or non-testing determined by factors other than educational need. To the extent that educational accountability and pupil-product evaluation are truly meaningful phrases, a planned, efficient and appropriate testing program is in order. It is true that

Table I

Age of Organized Programs

Years	% of Districts
1 - 2	8
3 - 5	18
6 - 10	35
11 - 15	17
16+	19
No answer	3

standardized testing data will usually tell less than is needed to assess educational effectiveness; however, it does tell more than would be known if no tests were given or no other measurements taken. Tests may be fairly criticized for a number of faults which they possess; however, it is less than sporting to criticize them for our mis- or non-use of the information they do contain.

Accordingly, four questions were put into the survey to assess the sources of operational authority and purpose as they presently exist. Table III presents the responses for the 433 districts of more than one building throughout the state to the question regarding lines of responsibility for their testing programs.

Table II

Percentage Distribution of Districts Without Organized Testing Program: What Characterizes Them?

Principal determines test policy and usage within own building	5%
Each teacher selects and uses standardized tests at his/her own discretion	1%
Tests are given only in cases of special need	1%
Each building independently establishes its own testing policy and/or committee	4%
Don't give any published standardized tests	1%

Table III

Percentage Distribution of "Who's In Charge?" of Testing Program

Testing program under one central testing committee and/or testing specialist	33%
Elementary school coordinator and/or committee—Secondary school coordinator and/or committee responsible for testing program	36%
Each building under its own testing committee and/or testing specialist	11%
Principal determines testing program of own building	10%
Does not apply (have only one building)	4%
Other arrangement	6%

It should be noted that about two-thirds of the districts reported a central source of authority: variously a testing specialist, curriculum coordinator, or the educationally ubiquitous committee. Blending this information together with that contained in the following tables, it is apparent where the source of much school testing authority lies. Directors of testing are rare and directors of curriculum apparently are not frequently found on testing committees. Thus, when a committee does exist it has little formal titular leadership. Accordingly, while in many districts committees exist, the locus of their authority appears to be somewhat outside of them. It is, however, encouraging that two-thirds of the districts at least possess the trappings of centrally coordinated authority.

Table IV

Membership of Testing Committees

Director of testing	6%
Assistant superintendent or superintendent	38%
Principal or assistant principal	81%
Teachers	44%
Counselors	75%
Director of curriculum	11%
Consultant, curriculum or guidance	30%

Table IV reports the distribution of offices amongst the districts (39 percent of total) who have an active testing committee.

Some equivocation may have existed over the use of the term "active" when inquiring into the presence of an "active testing committee" within the district. It seems probable to the writers that the implication of the findings is that 39 percent of the districts have a regularly assembling group, while 61 percent either do not have a committee or assemble one only when needed. It is not known how a committee with a more generalized function that includes testing in its purview answered this question.

Among the districts who reported the presence of an active testing committee, a high percentage (84%) reported that the committee functioned for the entire district (Table V). This value is higher than the authors' experiences reflect. Experience with many districts indicate there is a great deal of segmented planning throughout the state. However, the presence of a committee may be a force that will tend to encourage K-12 planning.

Table V

Percentage Distribution of the Scope of Testing Committees

Does committee function for entire district:		
	Yes	84%
	No, because:	16%
	There is a separate committee for elementary and secondary (14%), each building has a committee (17%), some other situation (17%)	

Finally in the management sector is the information concerning general involvement in the direction, review, and selection of tests. In this question respondents were asked to state who is involved in the three stages of program management and to rate the extent of this involvement on a 1-5 scale. The total number of usable responses came from 464 districts. The column headed "Direct-Evaluate" represents the personnel involved in the daily ongoing direction and/or continuing evaluation of the program. "Review" represents personnel that might be involved in initiation and carrying out of major program review. "Selection" represents personnel involved in selection of single tests, batteries of tests or groups of tests for use within the program. Table VI gives the weighted means for each of these three functions.

The horizontal stability of the values suggests that the three functions tend to fuse into a more generalized responsibility, i.e., those that direct also review and select. Regardless of the stated authority structure of the program, the actual responsibility for its operation very markedly rests with the counselors and principals. Teachers, the administration and organized committees enter as secondary agents in these functions.

Table VI

Weighted Mean Distribution of
Testing Program Involvement

Personnel	Function		
	Direct-Evaluate	Review	Selection
Teacher	1.8	0.7	1.9
Committee	1.2	1.4	1.5
Principals	0.6	2.9	3.0
Superintendent or assistant superintendent	1.9	1.9	1.9
Director of curriculum	0.8	0.8	0.8
Counselor or other pupil personnel specialist	3.6	3.1	3.7
Consultants from external educational agencies or services	0.7	0.7	0.8
Consultants from commercial test publishers	0.4	0.4	0.4

In response to the question about anticipated changes in testing programs in the next year, a surprising 46 percent of the districts indicated probable "significant or major" changes were anticipated.

To provide some definition of the anticipated changes, those districts (N = 251) were given nineteen options and were asked to indicate the nature of the expected change. The results displayed in Table VII cite the options and indicate the number of districts that either plan and/or need each, and those who do not feel it is necessary for them. It is significant, in light of the further findings of this survey, to note that those most needed or planned changes center around improvement in reporting and interpretation of test results.

Table VIIa

Frequency Distribution of Anticipated Testing Program Changes

	1 planned	2 needed	3 both	4 neither
1. To increase the use of reading readiness tests	36	40	19	156
2. To use a different reading readiness test than we are using	24	23	14	190
3. To increase the use of standardized reading tests (other than tests which are part of the instructional reading program materials	38	32	28	153
4. To use different reading tests than we are now using	28	26	23	174
5. To increase the use of individual intelligence tests	31	29	19	172
6. To increase the use of group intelligence or scholastic aptitude tests	38	22	20	171
7. To introduce or use a different group intelligence or scholastic aptitude test than we are now using	35	31	20	165
8. To introduce or use more multi-aptitude batteries	20	28	11	192
9. To introduce or use a different multi-aptitude battery than we are now using	22	12	4	213

Table VIIb

Frequency Distribution of Anticipated Testing Program Changes

		1 planned	2 needed	3 both	4 neither
10.	To increase the use of standardized achievement test batteries	37	31	30	153
11.	To use a different standardized achievement battery than we are now using	47	25	25	154
12.	To introduce or use more interest tests	20	42	25	164
13.	To introduce or use more personality or character tests	11	25	6	209
14.	To improve the scoring of tests	38	30	28	155
15.	To improve the methods of recording test results	37	41	36	137
16.	To improve the processing and reporting of test results to teachers, counselors, or administrators	50	64	55	82
17.	To develop more local (school district) norms	35	73	32	111
18.	To improve the interpretation of test results	53	69	57	72
19.	To improve the interpretation of test results to teachers, counselors, or administrators	57	74	65	55

The Dissemination of Test Information

One of the most frequent charges against those invested with control of testing information is the lack of dissemination of this information. A variety of reasons are offered for this, some doubtless of considerable merit. However, it is difficult to contest the uselessness of test data in a file. Frequently, our lack of confidence in the dissemination of data is directly related to our lack of adequately prepared teachers to ingest the information. Local districts must, then, either do the job themselves or face a situation that allows tests to be given while their information is not well understood.

To pursue the way in which test information is housed inquiry was made about the placement of data. As many responses as appropriate were requested, so the total well exceeds 100 percent. From Table VIII it may be seen that the pattern is consistent, with the counselor or principal being most involved with test data.

Table VIII

Percentage Distribution of Test Data Placement

In superintendent's or assistant superintendent's office	16%
In central office	20%
In principal's office	61%
In office of research director	3%
In office of testing director (if other than research director)	22%
With counselor or pupil personnel specialist	71%
With grade or homeroom teacher	35%
In pupil's cumulative folder	84%

Also of interest may be the numbers of copies of test data produced per administration, as gleaned from the responses reported in Table VIII. It is accepted that the original question might allow this interpretation.

Of greater interest, however, is the matter of eligibility to see the results. It must be remembered, that not all districts give "interest or vocational" tests and even fewer administer "personality" tests. It is advised that the worth of this question comes from the differences between persons for a given type of test rather than from comparing across tests.

Table IX

Copies of Test Data Produced
By School Districts

Number of Copies	Percent of Districts
1	11%
2	21%
3	32%
4	20%
5	10%
6	5%
7 and over	1%

Table X

Percentage Distribution of Persons Eligible to See Test Data

	Intelligence or Aptitude	Achievement	Person-ality	Interest or Vocational
Homeroom teacher (Sec.) or grade teacher (Elem.)	86%	87%	43%	70%
Any classroom teacher	84	87	40	74
Special teacher, speech, etc.	87	86	45	72
Principal	98	97	63	86
Chief school officers	70	70	41	62
Board of education	33	36	19	29
Guidance counselor	93	92	63	87
Specialist/consultant in health, psychology, etc.	83	81	54	72
No answer	–	–	31	31

On the assumption that a district has an established and systematized procedure for making test data available to its staff, the respondents were asked to indicate the one way in which data are made accessible. The responses here relate to jurisdiction and, eventually, to control of testing information. The figures (Table XI) suggest a fairly wide pattern of containment, although the "not available to teachers" is happily almost negligible.

Table XI

Percentage Distribution of Teacher Information Source

1. Test results are placed in the central office and any teacher who wishes may look them up.	31%
2. Test results are placed in files in the principal's office any teacher may see them in consultation with the principal.	17%
3. Test results are placed in files in the guidance counselor's office. Any teacher may see them in consultation with the counselor.	30%
4. Test results are sent directly to homeroom teacher who keeps them in his own file.	7%
5. Test results are confidential and are generally not available to teachers.	1%
6. Situation # 4, plus either # 2 or # 3 above	11%
7. Other	3%

A most common area of either uncertainty or conflict in education is the dissemination of standardized test information to parents and pupils as well as "in house" groups. These are the presumed consumers for whom the data was originally gathered.

The variety of ways in which information was regularly interpreted to pupils, parents, teachers, administrators and community groups was solicited. Table XII would seem to indicate the regular presentation of test information to the primary subjects and consumers, i.e., the children and their parents. Secondarily, teachers and administrators are presented with summary statistics. Perhaps, if true, this is how it should be.

One of the most expandable questions asked in the survey dealt with the ways in which test results are sometimes used. Respondents were asked to rate a list of possible uses on a 1 to 5 scale reflecting highest to lowest priority.

Table XII

Percentage Distribution of Methods for Interpreting Data

Written reports or profiles to pupils	53%
Written reports or profiles to parents	30%
Individual pupil conferences	85%
Individual parent conferences	66%
Group analysis with pupils	34%
Group analysis with parents	5%
Group analysis in community meetings	4%
Case studies in teachers' meetings	21%
Test analyses in teachers' meetings	33%
Consultant help in teachers' meetings	19%
Teachers' institutes	4%
Report of summary statistics to teachers	55%
Report of summary statistics to administrators	58%
Report of summary statistics to community groups	13%
Other	1%

This question will be studied in greater depth in another part of this report when the various use categories are reported in relationship to other questions.

Presumably the answers given by school districts to this question should be reflective of the purposes for which tests are given. The "why" should dictate the "what", "when" and "to whom" tests are administered. Some tests, under specific circumstances, are relevant to a particular function, e.g., curriculum diagnosis, while others are not. If curriculum evaluation is an avowed use of a test in a district then a test which does that should be used rather than one designed to be more diagnostic of individual students.

Table XIII reflects this question using weighted means in its first column. To retain some of the original data, the numbers of first choices for each category are also listed in parentheses following the weighted score.

Table XIII

Weighted Means and First Choice Responses to Test Result Use

	Weighted Mean	1st Choices
Teacher diagnosis of pupil strengths and weaknesses	3.4	(215)
Evaluation of curriculum	1.2	(10)
Development of education and vocational goals	1.8	(64)
Teacher analysis of class achievement	1.3	(12)
Placement in particular classes	1.7	(29)
Identification of the exceptional child	0.7	(8)
Determination of reasonable levels of achievement	1.9	(49)
Evaluation of education research	0.1	(0)
Development of parental understanding of pupil	0.7	(6)
Motivation for increased learning	0.5	(10)
Development of continuous program of teacher in-service education	0.1	(1)

The data seem to show that by far the most important use of test results is involved in the diagnostic relationship between teacher and pupil. Following at considerable distance is the broad spectrum of evaluation and development of educational goals and uses related to class and individual achievement.

Teacher/Staff Preparation in Test Data Use

Four alternatives were given the districts to describe the "provisions . . . . made by your system to assist teachers and other personnel to use test results most effectively". Twenty districts failed to answer this. Multiple choices were permitted. Table XIV shows the number of tallies per alternative.

No question was asked as to the districts' estimates of the effectiveness of these procedures. In few areas are we less adequately prepared to meet our

Table XIV

Methods Used for In-Service Education

Methods:	Frequency
General faculty meetings at least once each year devoted to testing program and interpreting test results	108
Building faculty meetings at least once each year devoted to testing program and interpreting test results	202
Departmental, grade, divisional or other sub-group faculty meetings at least once each year devoted to testing program and interpreting test results	165
In-service training facilities other than faculty meetings providing help in the testing program and test result interpretation	131

colleagues' and parents' questions, let alone our own. There is little doubt as to the nature of the answers if such a question had been posed.

Again twenty districts did not respond to the question of "who bears the primary responsibility for carrying out and directing such information and training meetings". The frequencies are used in Table XV below.

Table XV

Who Directs Testing In-Service Education

Title	Frequency
Director of testing	66
Director of curriculum (elem. or sec.)	21
Principal	129
Superintendent or assistant superintendent	39
Counselor or pupil personnel specialist	180
Other	10
No response	20

Note that the counselor and principal again carry this technical responsibility, their order reversed over the earlier questions relating to management and selection matters.

Chapter III

THE SECOND STUDY: TESTING SITUATIONS

While the first part of this report is concerned with testing programs and their implementation, the second part deals with the testing situations encountered. A "testing situation" is herein defined as occurring each time a test is given. Thus, if a district gives an achievement test in grades 3, 5 and 7 and an intelligence test in grades 4 and 6 that would constitute five testing situations.

To secure this information, every third questionnaire included an additional fourteen questions. These covered the tests given, how they are administered and scored, the results' subsequent availability, and the norms used for the tests.

The results reported here will be concerned only with reports of achievement and intelligence-aptitude testing practices in that 88 percent of the testing situations reported were of those types. Others, e.g., reading readiness, interest, personality, were not sufficiently used to permit generalization. The data in this part are based on 1417 testing situations, there being a range of one to thirty situations in respective districts, the mean being 12 per district. All data reported are in percentages of total per column.

The tables make two distinctions in their information. Elementary and secondary schools are separated between grades 6 and 7, although this distinction is becoming increasingly exceptionable with the advent of middle schools. Furthermore, information pertaining to achievement data is separated from intelligence and aptitude data, the last two being subsumed under the same category heading.

Table XVI reports the percentage distribution of testing situations in elementary and secondary grades for intelligence and aptitude, hereafter jointly called ability, tests and for achievement tests as related to the conditions of test administration. Eighty-one percent of the ability testing situations in elementary schools occur in classrooms, while only 41 percent of the ability tests are administered in that location in the secondary schools. The shift to group counseling is readily apparent for the higher grades where large groups share the limelight with classroom groupings.

Table XVII suggests that what is good for one is good for all. This may be of some interest to someone. Just how is this to be interpreted together with the information in the previous table? It would seem that what appears to be group counseling may in fact be merely administrative convenience.

Table XVI

Percentage Distribution of Testing Conditions

	Elementary Grades		Secondary Grades	
Test Situation	Ability	Achievement	Ability	Achievement
Classroom	81%	90%	41%	46%
Large groups	5%	6%	46%	39%
Small groups	5%	1%	9%	8%
Individually	5%	0	1%	0
Combination of Above	3%	, 3%	3%	7%
No response	1%	0	<1%	0

Table XVIII must be read carefully. It is designed to report responses to the question: "When a test is given, how frequently is it given?" Thus, a specific ability test is administered each year 95 percent of the time in secondary schools. It is used every other year in 1 percent of the cases at that grade level. This table does not say that 90 percent of the elementary youngsters are tested every year; rather, 90 percent of the ability tests are repeated annually.

Table XVII

Percentage Distribution of Pupils Tested

	Elementary Grades		Secondary Grades	
Proportion of Students	Ability	Achievement	Ability	Achievement
All	93	95	89	84
3/4 - 9/10	2	3	5	5
1/2 - 3/4	<1	1	1	1
1/4 - 1/2	<1	<1	1	3
1/4 less	0	<1	1	0
Only small no.	4	<1	1	6
Combination of above	0	<1	1	1
No response	1	0	1	0

Table XVIII

Percentage Distribution of Testing Situation Frequency

	Elementary Grades		Secondary Grades	
Frequency of Testing	Ability	Achievement	Ability	Achievement
Once each year	90	95	95	95
Twice each year	<1	2	<1	<1
More than twice each year	0	0	0	0
Every other year	1	1	1	<1
Some other regular schedule	4	0	1	1
Irregularly	3	<1	<1	1
No response	2	2	2	2

Table XIX contains information that may be extremely important *vis-á-vis* test use. Half of the ability tests are given in the fall, while achievement tests are more frequently given in the spring. This suggests that achievement tests are seen in most cases as summative rather than as prescriptive for educational planning. The time separation between ability and achievement testing may reflect a division of labor; however, it clearly restricts a district's ability to see achievement in relation to immediate ability.

Table XIX
Percentage Distribution of Test Administration

	Elementary Grades		Secondary Grades	
Time of Testing	Ability	Achievement	Ability	Achievement
Fall	50	31	51	37
Winter	14	10	15	11
Spring	28	54	30	45
Fall-Spring	0	2	1	0
Winter-Spring	2	<1	<1	<1
Fall-Winter	4	2	1	3
No specified time	1	1	1	2
No response	1	1	1	2

The shift in the responsibility for test administration is apparent in the next table. Counselors take over from classroom teachers in secondary schools and the principal's psychometric role almost disappears, particularly with ability tests. Of greatest interest, perhaps, is the difference in the teachers' role between the two kinds of tests. Clearly, the domain of intelligence-aptitude is seen as being less relevant to them than achievement.

In the scoring of standardized tests there is only a minimal distinction between ability and achievement tests. Rather, as shown in Table XXI, the difference occurs between the elementary and the secondary schools. There is a great variety of ways in which tests are scored. Clearly, the move is toward automated scoring processes and away from hand-scoring but the transition is gradual. Five years from now the trend should be much clearer. It is easy to

Table XX

Percentage Distribution of Test Administration:
Who Gives What?

	Elementary Grades		Secondary Grades	
Title	Ability	Achievement	Ability	Achievement
Classroom teacher	59	75	13	28
Guidance counselor	15	11	76	58
School psychologist	<1	<1	0	0
Consulting psychologist	0	0	<1	0
Principal or assistant principal	13	7	2	5
Superintendent	0	0	0	0
Other	4	<1	2	0
Combination	9	6	6	8
No response	0	0	1	<1

predict that the service agencies and local computer capabilities will be considerably increased by then and the laborious and often error-ridden hand-scoring operations will be less prevalent. Three quarters of the secondary schools already have access to automation. Note that achievement testing situations are somewhat more frequently machine-scored than ability tests.

Table XXII paints a mixed picture. The first four options all allow for the release of test information; however, the last three of these stipulate some qualifications. In the instance of secondary ability tests 85 percent of the results are "reported" under some conditions. It is even higher for achievement tests. Ability data are understandably more "confidential" than achievement data.

A final note: no attempt was made in the survey to specify what "test results" are (e.g., specific values, generalities, etc.), nor was the term "reported" further defined. Some variation may have existed in the minds of the respondents as they answered this question.

Table XXI

Percentage Distribution of Test Scoring

	Elementary Grades		Secondary Grades	
Test Scorer	Ability	Achievement	Ability	Achievement
Student	0	<1	<1	0
Clerk	5	5	4	3
Classroom teacher	32	34	3	10
Pupil personnel worker	11	3	16	6
Principal or administrator	5	2	1	2
Educational service organization	11	.14	30	24
Test publisher	22	32	33	45
Test scoring company	3	4	6	5
School owned scoring machine	4	3	3	4
Other	4	<1	1	0
No response	3	2	2	<1

Table XXII

Percentage Distribution of Test Situations Reported to Children

	Elementary Grades		Secondary Grades	
Are Tests Reported	Ability	Achievement	Ability	Achievement
Yes, routinely	3	8	35	41
Yes, some cases	5	13	7	15
No, explanation routinely	13	14	18	13
No, explanation some cases	34	37	25	22
No, test confidential	39	21	12	7
Combination	2	<1	<1	1
No response	4	6	2	1

Multiple responses were sought regarding the uses to which test results are put. Table XXIII shows some interesting features.

"Ability grouping", in spite of years of systematic research demonstrating its general futility, remains vigorously present. Why do we persist in reporting research when data are as ignored as these?

The relative weakness of use for "grading students" is comforting with respect to achievement tests. This mis-use of testing information may be dying slowly, though secondary schools continue to show some persistence. How ability tests can be used to grade students in 5 percent of the cases is difficult to understand. The use of ability tests to evaluate "curriculum" and "teaching" defies reaction.

To use these instruments to counsel students and parents, to diagnose learning difficulties and, with achievement tests, to evaluate curriculum (not teaching) are all frequently reported and are appropriate uses.

The 1 percent who report "not used" for elementary achievement are thanked for their frankness. One suspects they may have more colleagues than are acknowledged.

Table XXIII

Percentage Distribution of Test Uses

	Elementary Grades		Secondary Grades	
Test Usage	Ability	Achievement	Ability	Achievement
Ability grouping	34	43	25	34
Counseling students	37	39	82	70
Grading students	5	8	6	13
Evaluate curriculum	27	54	25	51
Evaluate teaching	9	25	7	19
Diagnosis of learning difficulties	76	73	60	63
Counseling parents	50	47	48	48
Other	2	2	1	2
Results not used	0	<1	0	0

Table XXIV presents some unusual response characteristics. Nearly a quarter of the respondents did not indicate the single "most important use" of test data. Among those who did, however, the data indicated that both achievement and ability tests are used to diagnose learning difficulty in the elementary schools while the predominant use in the secondary grades is for counseling students.

Table XXIV

Percentage Distribution of Most Important Use Data

	Elementary Grades		Secondary Grades	
Use	Ability	Achievement	Ability	Achievement
Ability group	12	11	6	7
Counseling students	10	7	48	41
Grading students	0	1	1	1
Evaluate curriculum	3	11	1	8
Evaluate teaching	1	3	1	2
Diagnosis of learning difficulties	45	38	17	17
Counseling parents	7	5	1	1
Other	1	1	1	2
No response	21	23	24	21

It should be noted that in Table XXV the "publishers" norms are almost certainly national in scope and should be so subsumed. Further, nearly one quarter of the testing situations apparently apply more than one norm reference point. This is encouraging!

Also encouraging is a fairly substantial tendency to report test data via more than one statistical language. However, as is shown in the options of Table XXVI, the questionnaire language was confusing; all are standard scores (except profiles), not just the 7 to 23 percent so recorded.

Clearly, the I.Q. score is used to report ability measures in elementary schools, whereas percentiles are used for achievement tests. The data suggests use of the more manageable scores such as bands or stanines may be on the increase. It is certain, whatever the present trends, the grade equivalent score reflects a day when only less understandable test language was available.

Table XXV

Percentage Distribution of Norms Used

	Elementary Grades		Secondary Grades	
Norms	Ability	Achievement	Ability	Achievement
Local	16	24	19	27
Regional	4	6	1	2
National	92	91	92	93
Publishers	<1	1	1	2
Other	<1	0	0	0

Table XXVI

Percentage Distribution of Test Language

	Elementary Grades		Secondary Grades	
Test Language	Ability	Achievement	Ability	Achievement
IQ score	83	6	48	9
Stanine	7	22	5	16
Standard scores	7	15	15	23
Grade equivalents	9	73	13	49
Age equivalents	11	7	8	8
Percentile rank	30	51	60	71
Band scores	2	6	10	7
Profiles	0	0	2	0
Other	0	0	0	0

Table XXVII is easily read and offers little news. One should be pleased that only three testing situations surveyed responded that tests were too confidential for teachers to see. This is consistent with the results reported earlier in Table XI.

Table XXVII

Percentage Distribution of Test Availability to Teachers

	Elementary Grades		Secondary Grades	
Test Availability	Ability	Achievement	Ability	Achievement
Yes, in classroom files	51	56	11	24
Yes, in central files	32	28	53	50
Yes, in consultation with principal, etc.	7	9	23	19
No, tests confidential	0	0	1	1
Combination	7	5	10	6
No response	3	2	3	1

Table XXVIII offers a wide range of responses as could be anticipated. In excess of 85 percent of the testing situations generate data that may be made available to parents under some conditions. It seems that only the elementary school ability data are withheld with any degree of frequency.

Table XXVIII

Percentage Distribution of Test Availability to Parents

	Elementary Grades		Secondary Grades	
Test Availability	Ability	Achievement	Ability	Achievement
Yes, routinely	2	12	12	15
Yes, parents request, school approves	13	21	18	31
No, explanation routinely	12	13	8	8
No, explanation parents request, school approves	55	47	51	38
No, test results confidential	11	3	5	3
Combination	3	2	4	5
No response	4	2	2	0

As was noted with an earlier question, the terms "available" and "results" are not specific and may reflect a fairly wide range of practices.

Perhaps the greatest general statement that can be made is that the data can be had for the asking. But one must know who to ask, what to ask for and feel confident enough to ask. One wonders how many ask.

Table XXIX again reflects the management of test data; reporting test scores to parents and children. Classroom teachers, principals, specialists and combinations thereof carry out this function in the elementary school only to be heavily replaced in the secondary schools by the counselors.

Table XXIX

Percentage Distribution of Who Reports Scores to Children/Parents

	Elementary Grades		Secondary Grades	
Title	Ability	Achievement	Ability	Achievement
Classroom teacher	37	53	3	7
Guidance counselor	11	9	74	70
School psychologist	0	<1	0	0
Principal	12	8	3	5
Other	1	<1	0	0
Not interpreted to parents or child	7	2	3	3
Combination	30	25	14	16
No response	3	2	3	0

Chapter IV

TESTING PATTERNS

Part of the survey inquired into the testing behavior of districts with respect to particular instruments. It is with mixed feelings that the authors report this information. Clearly, popularity of use is hardly a sufficient criterion for test selection, particularly with tests of scholastic achievement. But the question, "What do other districts give?" is too frequently asked to permit anything other than reasonably complete reporting of this information.

Frequencies are reported in two ways. The number of districts using each test is followed by the number of testing situations in which each is used. Accordingly, if a district gives a certain aptitude test twice the tally would be one for districts and two for testing situations. It is hoped the latter will be helpful in placing this information in better perspective.

These data are based on the one-third sample population described earlier. The basic number of districts here is 131. Only tallies for group tests are reported. While some districts listed the Wechsler and others, the intent of the survey was to secure group test use only.

Clearly some patterns exist. Intelligence tests, used in the lower grades more than in the upper, tend to occur two or three times in a child's experience. The most typical pattern (not observable from Tables XXX *a* and *b* but elsewhere) is two elementary administrations and one secondary, usually in the junior high. If only two are given, one is given in elementary and about half the others are given again in elementary and half in the secondary.

Table XXXa

Frequency Distribution of Test Usage

Name of Test	Districts Using	Testing Situations
General Intelligence Tests		
Calif. Test Mental Maturity - Long Form	7	10
Calif. Test Mental Maturity - Short Form	83	229
Henmon-Nelson	4	8
Kuhlmann-Anderson	11	17
Lorge-Thorndike	52	127
Otis	17	30
Otis-Lennon	20	38
School & College Ability Test (SCAT)	17	24
SRA TEA	9	10
(Scholastic) Aptitude		
Academic Promise Test	7	7
Differential Aptitude Test	98	116
Iowa Algebra Aptitude	22	24
Orleans Algebra Aptitude	3	3
SRA PMA	5	7
Reading Readiness		
ABC Inventory	17	18
Gates	3	3
Harrison-Stroud	4	6
Lee-Clark	6	7
Metropolitan	62	70

Table XXXb

Frequency Distribution of Test Usage

Name of Test	Districts Using	Testing Situations
Achievement Batteries		
California	38	113
Cooperative	4	10
Iowa Tests Basic Skills	40	113
Iowa Educational Development	32	65
Metropolitan	28	80
SRA	13	46
STEP	9	12
Reading Tests		
Gates Primary	3	4
Gates MacGinite	16	35
Scott Foresman Basic	8	10
Inventories		
Kuder Preference	31	32
Strong Vocational Interest	8	9

(All other tests mentioned were given fewer than five times by two or fewer districts only. These criteria hold throughout this tabulation.)

Scholastic aptitude tests are limited almost exclusively to the secondary schools. The *Differential Aptitude Tests* seem to be the only test repeated. The authors feel obliged to observe an old testing canon to the effect that putting all one's evidence in one test, and not repeating it to assess stability, has dangers. It is not sufficient to say that these are reliable instruments; of course they are, but they are more reliable for groups of youngsters than for individuals.

The running mode for achievement batteries seems to be between three and four administrations. (There are those districts that use more than one brand of intelligence or more than one brand of achievement tests. While they

limit generalizations, they do not occur frequently.) Generally, two administrations occur at elementary and one or two in secondary schools.

Leaving the table but continuing on the matter of particular test popularity, the authors' principle reason for feelings of reservation about printing this kind of information should also be stated.

Particularly in the area of achievement tests, the spawning of new forms of old tests is an increasing phenomenon. As such, whether a test is "popular" or not at a given point in time depends as much on its recency, and the recency of its competition, as it does upon the quality of the instrument. The availability of a test as a "recent" event may be illustrated by the fact that three major achievement tests will produce new editions in the twelve months following this writing (summer, 1970). What the popularity of tests, both old and new, will be a year from today is, of course, uncertain.

Tests, particularly achievement tests, should be selected on a basis other than what the other school is doing. The competencies of districts to make insightful decisions are about uniformly distributed and each must wrestle with the same basic questions in test selection. Other than practical concerns of cost, readability and flexibility of out-put, time to administer and servicing, the sole criterion of test selection is curricular fit. This emphasizes the authors' distress at the relative lack of participation by curriculum leaders in test selection. Only by determining which instrument most closely asks questions relevant to the local curriculum with appropriate grade level expectations can an achievement test render believable results.

Chapter V

TEST INFORMATION USAGE

Education is frequently charged with being badly hung up on traditionalism and self-perpetuation. Certainly many of our practices are open to question; it is perhaps a hopeful sign that some of our less defensible practices are presently undergoing modification.

Among our many rather *pro forma* acts is the annual exercise for many youngsters of taking standardized tests, the results of which are used less than they should be if tests were to be cost- or time-justified. Used or not, testing programs are "good to have", or so our actions would seem to imply.

In the interests of assaying school districts' declared uses of test information, a question was inserted into the survey instrument that offered a variety of possible uses. This chapter of the report will address itself to those declared uses in relation to the care-storage-control of the information, with respect to how the information is disseminated and as to how teachers are aided to better understand and interpret test information.

The question on which this chapter is based asked districts to rate, on a 1 to 5 scale, the "ways in which test results are sometimes used". Thirteen options were provided. Reviewing the responses enabled the authors to say that the districts tended to group themselves into four types: those who use results to assess individual achievement (Type I); those concerned with test usage for motivational purposes (Type M); those who emphasize the research and development aspects (Type R & D); and a group almost uniformly equal in their emphases in the three categories (Type E). A sub-sample was drawn from the total population of those districts that most clearly represented each of the four types. Nearly 40 percent of the districts in the state were included in this sample, with the designations being judgmental.

Individual Achievement	Type I	N = 56
Motivation	Type M	N = 51
Research & Development	Type R & D	N = 46
Balanced Equivalent	Type E	N = 56

The question was then asked as to how these four types of districts responded to three of the other questions in the survey. The data follow.

Where Are Tests Kept?

There were originally five alternatives identified in addition to an "other" category. One ("tests are too confidential and are not available to teachers") received so few responses it was happily discarded.

Table XXXI

Percentage Distribution of Test Placement
by Type of School District

	Type of District by Test Use			
Where tests are placed	I	M	R&D	E
1) Placed in file in central office - teachers may look up	34	37	24	27
2) Placed in principal's office; seen via principal thru consultation	13	37	41	41
3) Placed in counselor's file; seen via counselor thru consultation	30	16	21	18
4) Kept in the classroom	13	10	14	14
5) Other	10	0	0	0

Table XXXI presents a 4 x 5 comparison, the "4" being the four types of school districts identified and the "5" being the five possible choices in response, including "other". The values reported are percentages of the vertical, type-of-district, column.

To read this table note that 34 percent of the Type I districts place the data in the central office and permit direct access by teachers. Thirteen percent of those districts keep the data in the principal's office and 13 percent keep them in the classroom. Some 30 percent of that district-type use the counselor's file. The distributions for each district-type may be read in the same way.

There are statistical differences. When the data were tested for differences they yielded a chi square value of 23.47, which is just short of statistical significance at the .05 level. The differences in the table, then, are not statistically significant by a conventional standard of value. However, the size of the chi square suggests that the differences are notable if we take something slightly less than the .05 confidence level. It is safe to say the obtained differences in the table exceed occurrence by chance alone 93 times out of 100.

Accordingly, there appears to be some tentative kind of relationship between district-type and test data placement. The "R & D" and "E" type districts appear to be monotonic in their placement of test data and that in the principal's office. This would seem to be in keeping with the stated uses of these data. In contrast the other two types of districts are dichotomous in their test data storage. In both cases records are situated in the central office, perhaps as a basic repository. Those districts primarily inclined toward individual assessment also tend to place test results in the counselor's office where they are available for individual reference. The "Motivation" districts tend, in contrast, to favor the principal's office.

How Are Results Reported-Interpreted?

Again the filter is the district-type. The essential question concerns whether districts which tend to use test data differently (hence, have different purposes) also reflect differences in other testing program attributes. The survey question asked respondents to indicate the methods used regularly in interpreting tests to pupils, parents, teachers, administrators and community groups. Thirteen of the methods received sufficient tallies to warrant reporting and analysis. The data in Table XXXII report the percentage of *yes* responses. Each category-response possesses its own chi square because the comparison here was *yes–no.* Each such response category totals 100 percent. Thus, there were 53 percent *no* responses to "written reports or profiles to pupils" in the "I" column against the 47 percent reported.

Table XXXII

Percentage Distribution of 'Yes' Responses Regarding Test Interpretation Methods

Methods used to report test data	Type of District by Test Use I	M	R&D	E	Chi-square value* *Yes* vs. *No*
a. Written reports or profiles to pupils	47	38	64	59	2.50
b. Written reports or profiles to parents	29	50	78	33	8.13*
c. Individual pupil conferences	87	86	52	79	14.85*
d. Individual parent conferences	76	33	50	100	3.67
e. Group analysis with pupils	42	100	0	75	2.90
f. Group analysis with parents	7	38	20	33	7.97*
g. Group analysis with community meetings	4	44	40	75	35.17*
h. Case studies in teachers' meetings	25	55	38	34	7.36
i. Test analyses in teachers' meetings	29	57	45	43	4.44
j. Teachers' institutes	4	78	71	94	86.88*
k. Statistical summaries to teachers	53	67	33	33	1.10
l. Statistical summaries to administrators	45	50	0	33	0.19
m. Statistical summaries to community groups	9	48	58	44	29.65*

*chi square significant at .05 level of confidence or greater

There are, then, six statistically significant differences shown in the table among the pairings. The implied 2 x 4 chi squares show that, most notably, the districts that stress the individual assessment function most frequently reject certain functions for testing data. On the contrary, statistical summaries, profiles to pupils and parent conferences are not related to type of district or to pronounced tendency.

The chi square statistic does not enable one to determine the source of the difference within each 2 x 4 table. As has been done in the past, the reader is invited to seek out his own interpretations of the spreads. Exercising that privilege, the authors note that the Type E districts appear markedly different on most dimensions from the others and that the Type I districts in many cases establish their own pattern. (Note *d, f, g, j* and *m* for the latter). The four occurrences of total deviation (100% or 0) surprise only by their totality.

Certain methods of test reporting, then, appear to bear a relationship, in their popularity, to certain kinds of districts. For some functions there appear to be philosophical purposes, not spontaneous operations as has been charged.

Assisting Teachers to Use Results

It is sometimes viewed with irony that districts spend rather generous amounts of time and money to secure test data and then invest little or nothing to aid their teachers in the knowledge necessary to good use. Again, the question under consideration is the relationship between district-type and the attendant endeavors to encourage teacher use and understanding.

The survey question asked for the provisions made by the system to assist teachers and other personnel to use tests more effectively. Four choices were offered. The format of Table XXXIII and its reading is the same as that of the previous table.

Respectable chi squares were attained in three of the four categories but only the provision of "at least annual department, grade or divisional meetings" attained statistical significance. There appears then, to be only a modest relationship between teacher in-service practice and district-type.

Table XXXIII

Percentage Distribution of 'Yes' Responses Regarding Provisions for Test Use

Provisions to use test data	Type of District by Test Use				
	I	M	R&D	E	Chi-square value* Yes vs. No
General, at least annual, faculty meetings for test data	27	38	14	33	1.13
Building, at least annual, faculty meetings for test data	54	58	0	0	7.38
Department, grade or divisional meetings, at least annual	25	63	50	63	8.41*
In-service training other than faculty meetings	29	67	25	40	5.65

*chi square significant at .05 level of confidence or greater

Chapter VI

TEST USAGE: DISTRICT CHARACTERISTICS

An additional way of looking at the use of test information, while essentially a non-dynamic one, is the potentially interesting act of relating test use to some of the physical attributes of a district. Speculation may be made on the relationship of use characteristics to the size of a district and to the political-geographical climate. Further, when these two characteristics are compounded one has the increased capability of seeing test use set in a relatively meaningful framework.

The base data in this section of the report derive from the questions in the survey that inquired into the population of the districts and into their *character.* The latter offered distinctions between rural, urban, suburban, metropolitan and some possible combinations of these. Analysis of the data suggested three population categories and four *character* divisions. This resulted in a 3 x 4 matrix, only two of whose cells were not penetrated. Graphically, the chart below presents the lay-out.

Chart I

Compound District-type, With Identifying Number Code

Population Served	Rural	Rural-Urban	Urban-Metropolitan	Suburban
5,000 or less	Type 1	Type 2	--	Type 3
5,000 - 25,000	Type 4	Type 5	Type 6	Type 7
25,000 or more	--	Type 8	Type 9	Type 10

While any division of the districts into a matrix-format is open to question, the one adopted provided a modest number of *zero* cells (2) and left the rest with a fairly uniform distribution. The present system shows there to be no *rural* districts of greater than 25,000 population, which is highly believable, nor are there any *urban-metropolitan* districts of less than 5,000, also not too unreasonable.

Chart II shows the number of districts that fell into each category. The ten types of districts thus identified were used to assay the responses to the four *test use* questions in the survey. Each will be discussed separately below *vis-á-vis* the classification system.

Reported are chi square significance levels for the values in the tables for the compound system as well as its components, labeled "size" and "character".

Chart II

Frequency of District Type

Population Served	Rural	Rural-Urban	Urban Metropolitan	Suburban
Under 5,000	152 (I)	48 (III)	—	7 (III)
5,000 - 25,000	34 (IV)	88 (V)	10 (VI)	49 (VII)
25,000 or more	–	8 (VIII)	14 (IX)	26 (X)

District-Type Related to Declared Use

The relation of the physical attributes of school districts to the ways in which they report the use of the data is shown in Table XXXIV.

Only two modest tendencies manifest themselves in the table. There appears to be a statistically significant relationship between the character of the district and the first two options. With increased urbanization of a district there is a tendency to use the data as shown. However, the key element, the compounding of the two attributes as detailed in the preceding charts, produces essentially nothing. Test use, then, when defined by the present options, appears to be unrelated to district size or character.

Table XXXIV

Significance of Chi Squares Between
Characteristics of Districts and Declared Uses

Ways Test Results Are Used	Variable I District Size	Variable II District Character	Compounded District-type
Teacher diagnosis of pupil strengths and weaknesses	N.S. (not significant)	.01	N.S.
Evaluation of curriculum	N.S.	.01	N.S.
Develop educational goals	N.S.	N.S.	N.S.
Teacher analysis of class achievement	.10	N.S.	N.S.
Class placement	N.S.	N.S.	N.S.
Identification of exceptionals	N.S.	N.S.	N.S.
To determine reasonable levels of achievement	N.S.	N.S.	N.S.
Evaluate Educational research	N.S.	N.S.	N.S.
Develop parent understanding of child	.10	N.S.	N.S.
Motivate increased learning	N.S.	N.S.	N.S.
Develop in-service program for teachers	N.S.	N.S.	N.S.

District-Type Related to Data Placement

A more productive analysis of the survey data related the district characteristics to the placement of test information in the buildings. While of itself the physical location of test scores may seem unimportant, experience suggests that access has to do with responsibility and control. Data must be visible and available to its potential user if he is to be a frequent consumer.

The table again reports the chi square value of the 6 x 19 matrix. Respondents were asked to select the one most correct response, so a choice-by-choice selection or rejection cannot be used with these data as they are in the next two tables. The vertical dimension describes the options, the horizontal describes the ten district-types. The cell values are the percentage of districts of each type indicating any choice.

The chi square value of the table data is statistically significant at the .01 level of confidence (97.23). Collapsing the categories did not markedly improve the statistic here or throughout the rest of this analysis.

Suburban districts of middle size are proportionately much more counselor-bound than others and clearly reject the central office and principal's office as a source of test information.

Other interpretations are left to the reader. Again, the chi square statistic does not permit localization of the trends away from statistical norm.

District-Type Related to Interpretation to Public

The survey instrument inquired into the methods used to report test data to parents, teachers, pupils, administrators, etc. These responses have been related to the individual and compound district characteristics. The data in the table follow preceding formats, including the district size and district character categories. The cells report the significance levels of each situation. Because the stimulus-question asked for as many choices as were appropriate, the chi-squares have been computed on each response category, the choice being *yes* or *no* for each.

Jumping out from the page are the highly significant choices centering around reporting summary statistics to teachers, administrators and community groups. Though the initial language is different, the "test analyses for teachers" is, in retrospect, perhaps of the same *genre* as the "summary statistics" for different groups. Clearly, there is a relationship between this use for test data and district characteristic.

Table XXXV

Percentage Distribution of District-Type vs. Test Data Placement

Where Results Are Kept	District-Type I	II	III	IV	V	VI	VII	VIII	IX	X	Total
Placed in central office files and available to any teacher	43	44	14	38	30	10	12	13	0	15	22
Placed in principal's office and available through the principal	19	6	29	18	17	10	12	37	22	19	16
Placed in counselor's office and available through the counselor	26	40	29	35	27	30	53	0	14	27	31
Kept in the homeroom	4	6	0	0	10	20	6	25	21	20	8
Data confidential; not available generally to teachers	2	2	0	0	0	0	0	2	0	0	1
Other	6	2	8	9	16	30	15	25	43	19	12
Total	100	100	100	100	100	100	100	100	100	100	100

Significant differences, though not as sharp, are found also with respect to reporting data to parents and community groups. They do not occur when the focus is on the child. Accordingly, the quasi-administrative practices of classifying children are associated with the physical attributes of a district. Additionally, individual parent conferences, but not pupil conferences, are related to district size and character.

Table XXXVI

Significance of Chi Squares Between Characteristics and Interpretation Methods

Methods used to interpret tests	Variable I District size	Variable II District Character	Compound District-Type
Written reports to pupils	N.S.	N.S.	N.S.
Written reports to parents	.10	.05	.10
Individual pupil conferences	N.S.	N.S.	N.S.
Individual parent conferences	.001	.05	.01
Group analysis with pupils	N.S.	N.S.	.05
Group analysis with parents	N.S.	N.S.	N.S.
Group analysis in community	.005	.05	.10
Teacher's case studies	.10	N.S.	N.S.
Test analyses for teachers	.001	.001	.001
Teacher institutes	.05	N.S.	.01
Report of summary statistics to teachers	.05	.05	.01
Report of summary statistics to administrators	.001	.001	.001
Report of summary statistics to community groups	.001	.001	.001

In the matter of using data to understand student growth and report it the distribution of the responses was sufficiently random so that differences could have occurred by chance alone.

District-Type Related to Improved Use

Inquiry into the in-service training of teachers to better utilize the information contained in standardized test data produced a question that assessed school practices on the same dimensions as shown in Table XXXVI.

The question asked for appropriate descriptions of means used to "assist teachers . . . to use test results". Accordingly, the four choices are presented on a *yes-no* basis providing individual chi square analyses for each. The data reported in Table XXXVII are significance levels for each category.

Table XXXVII

Significance of Chi Square Between Characteristics and In-Service Provisions

Provisions used for in-service training	Variable I District Size	Variable II District Character	Compound District-Type
At least annual general faculty meetings devoted to test interpretation	.005	N.S.	.10
At least annual building faculty meetings devoted to test interpretation	.001	.001	.001
At least annual depart-mental grade, divisional meeting for test inter-pretation	.001	001	.001
In-service activities other than faculty meetings for test interpretation	.005	.001	.01

The implications of these data are broader than they at first seem. The tone, control, and specificity of a district-wide meeting are different than, say, a grade level meeting, regardless of size. "General" faculty meetings can only be general in their focus.

The table shows clearly significant differences between the categories. The apparently universal phenomenon of faculty meetings seems to occur in a non-predictable pattern when character is concerned but is highly dependent on district size.

All other differences are significant indicating a relationship between the size or type of district and the techniques employed to assist teachers and other personnel to better understand test information. The table, however, does not indicate the direction taken by this relationship. If other findings of this study are to be taken into account it might be necessary to conclude that the direction of the in-service efforts are anything but positive.

CONCLUSION

In this report the authors have attempted to show, in addition to the standard description of testing practices throughout the state in 1970, that the ways in which test data are utilized may be related to other factors in the district. It has been demonstrated, for example, that districts of similar dimensions behave differently than dissimilar districts with respect to controlling test information, in-service training of teachers for test use, and reporting data to the community. In other areas there appear to be no differences.

An increasing proportion of districts in Michigan report structured testing programs with organized committees assuming the greatest role in the management of these activities. Counselors are seen as bearing the primary responsibility in secondary schools for operation of program, dissemination and interpretation of results, and education of potential test users.

When taken as single testing situations there is, even yet, little particularization of tests to the specific needs of a student; rather the trend continues to be, "What's good for one is good for all".

One view of the data suggests that test results are fully disseminated to child and parent alike. However, another perspective suggests the more tempered view that it is still difficult to ascribe adequate usage and understanding to the information derived from testing programs.

Implanted in much of this report is, of course, a series of values held jointly by the authors, having to do with the purposes and uses to which test data are put. The articulation of these values has been a means of their expression: more use should be given to test data, which can only be accomplished by creating more literate consumers. The reported activities of

Michigan school districts suggest that in certain places the locks may be coming off the files so that the information is not the sole domain of a select few.

This survey has allowed the Michigan school districts to state their testing policies and has attempted to interrelate the various elements of these practices. The time is ripe for a study to investigate relationships, if any exist, between these test practices and the quality of the district when measured against some external criteria.

There is need for an indepth study to determine if the quality of a testing program makes any impact on the education within the district. In other words, does the testing program really assist the teachers to do a better job of teaching; or the counsellors to do a more effective job of advising students; does all this really make a difference in the final product, the student? After all, isn't that what education is all about?